A Pragmatic Approach for Facial Expression Analysis using Machine Learning Algorithms.

Synopsis of the Research Proposal Submitted for the Registration in the

Ph.D. Program (Part-Time) in Engineering

Vishnu Kumar Barodiya

Under Supervision of

Dr. A. K. Saini

Dean of IcfaiTech, Jaipur

THE ICFAI UNIVERSITY JAIPUR

Introduction:

Facial expression analysis is a field of study that focuses on interpreting and understanding human emotions and intentions through the observation and analysis of facial expressions. It involves the application of various techniques, including computer vision, machine learning, and psychology, to recognize and classify facial expressions into distinct emotional states.

Facial expressions are a fundamental means of nonverbal communication, allowing individuals to convey their emotions, attitudes, and intentions. They provide valuable cues that can be used to infer a person's mental and emotional state, helping us understand their feelings, reactions, and intentions better. Facial expression analysis aims to capture, interpret, and quantify these expressions to gain insights into human behavior and improve various applications such as human-computer interaction, psychology research, market research, and more.

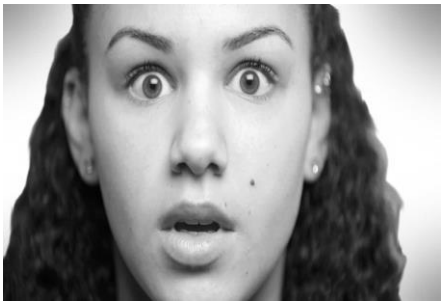
Advancements in computer vision and machine learning techniques have revolutionized the field of facial expression analysis. Facial recognition algorithms can detect and track facial features, such as eyes, nose, mouth, and eyebrows, in real-time or from images and videos. Machine learning models can be trained on large datasets of labeled facial expressions to recognize and classify different emotions accurately. These models can detect subtle changes in facial muscle movements and patterns to infer emotional states like happiness, sadness, anger, surprise, fear, and disgust.

We mainly focused on generating facial expression for seven emotions: happy, sad, fear, surprise, anger, disgust, and contempt.

Anger

Facial movements: Eyebrows pulled down, upper eyelids pulled up, lower eyelids pulled up, middle lips rolled in, and lips may be tightened.

Fear



Facial movements: Eyebrows pulled up and together, upper eyelids pulled up, mouth open.

Disgust

Facial movements: Eyebrows pulled down, nose wrinkled, upper lip pulled up, lips loose.

Happiness

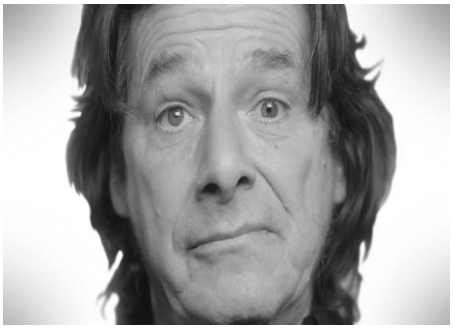
Facial movements: Muscle around the eyes tightened, "crow's feet" wrinkles around the eyes, cheeks raised, lip corners raised diagonally.

Sadness

Facial movements: Inner corners of eyebrows raised, eyelids loose, lip corners pulled down.

Surprise

Facial movements: Entire eyebrow pulled up, eyelids pulled up, mouth hangs open, pupils dilated.



Contempt

Facial movements: Eyes neutral with the lip corner pulled up and back on one side.

BASIC TERMINOLOGIES

1. Face Detection:

Face detection is a computer vision technology that involves automatically identifying and locating human faces within digital images or video frames. It is a fundamental task in many applications, including facial recognition, biometrics, surveillance systems, virtual reality, and social media.

The primary goal of face detection is to accurately detect the presence and position of faces in an image or video stream. This process involves analyzing the visual patterns and features that are characteristic of human faces, such as the arrangement of eyes, nose, mouth, and other facial landmarks.

In recent years, significant advancements in face detection have been made, primarily driven by deep learning techniques, particularly convolutional neural networks (CNNs). CNNs are capable of learning complex patterns and features from large amounts of training data, enabling them to achieve remarkable accuracy in face detection.

The face detection process typically involves several steps. First, the input image is analyzed to identify regions of interest that may contain faces. This step, often referred to as candidate region selection, can involve various techniques such as sliding window approaches or more advanced methods like region proposal networks (RPNs).

Once potential face regions are identified, the next step involves evaluating these regions to determine whether they actually contain faces. This is done by applying a face classifier, which is trained to distinguish between face and non-face regions. The classifier may utilize machine learning algorithms, such as support vector machines (SVMs), decision trees, or neural networks.

To improve detection accuracy, additional techniques can be employed, such as multi-scale analysis, which involves detecting faces at different scales to handle variations in size and perspective. Additionally, algorithms can utilize facial landmarks or features, such as the eyes, nose, and mouth, to refine the detection results and provide more detailed information about the detected faces.

2. Facial Feature Extraction:

Feature extraction converts a picture of the face into a higher-level representation of the shape, color, texture, and spatial configuration of the face or its components. Feature extraction will reduce the picture dimensions while keeping the important information. Feature extraction is important in formulating better emotion categorization as the extracted facial features are inputs to the classification module, which finally categories different emotions.

Facial feature extraction is a crucial step in computer vision and facial analysis tasks, including facial expression recognition, face recognition, and facial attribute detection. It involves detecting and localizing specific facial landmarks and regions of interest within an image or video frame.

The goal of facial feature extraction is to identify and locate key facial components, such as the eyes, eyebrows, nose, mouth, and chin, as well as other relevant features like facial contours and landmarks. These features provide valuable information for subsequent analysis and understanding of facial expressions, emotions, and identity.

3. Expressions Classification:

Facial expression, or emotion expression, is one or more motions or positions of the face muscles. There are various classification methods used to extract expressions. These

movements express the emotional state of the person to observers. It is a form of non-verbal communication. It plays a communicative role in interpersonal relations.

Expressions classification refers to the process of categorizing facial expressions into different emotional states or categories. It is a fundamental task in facial expression analysis and plays a crucial role in understanding human emotions and behavior.

Human beings communicate and express their emotions through a wide range of facial expressions, including happiness, sadness, anger, surprise, fear, disgust, and neutral expressions. Expressions classification aims to automatically recognize and classify these emotional states based on the analysis of facial features and movements.

Advancements in computer vision and machine learning techniques, particularly deep learning, have greatly improved the accuracy and effectiveness of expressions classification. These techniques involve training models on large datasets of labeled facial expression images or videos to learn patterns and features associated with different emotional states.

Expressions classification typically involves several steps. First, facial detection algorithms are used to locate and isolate the face within an image or video frame. Next, facial landmarks or key points, such as the position of the eyes, nose, and mouth, are identified and tracked. These landmarks provide valuable information about facial movements and expressions.

Fig. 1 Basic terminology of face expression recognition

DIFFERENCE:

FACE RECOGNITION AND FACIAL EXPRESSION RECOGNITION

1. **Face Recognition:** Face recognition is the process of identifying or verifying an individual's identity based on their facial features. It involves comparing the facial characteristics of a person in an image or video frame against a pre-existing database of known faces. The primary objective of face recognition is to determine whether a face belongs to a specific individual or to find a match in a database.

   Face recognition algorithms focus on extracting unique facial features, such as the shape of the eyes, nose, mouth, and overall facial structure. These features are then compared to a set of reference features stored in a database to determine a match. Face recognition technology has applications in security systems, access control, surveillance, and personalized user experiences.

2. **Facial Expression Recognition:** Facial expression recognition, on the other hand, aims to understand and interpret the emotional states conveyed by facial movements and expressions. It involves analyzing the changes in facial features, such as eyebrow position, eye openness, mouth shape, and overall muscle movements, to infer the underlying emotions being expressed.

   Facial expression recognition algorithms detect and track facial landmarks and analyze their configuration and changes over time. These algorithms typically utilize machine learning techniques to classify facial expressions into categories such as happiness, sadness, anger, surprise, and more. The applications of facial expression recognition range from emotion detection in human-computer interaction systems to psychological research and affective computing.

2. Literature Review

Deepak Kumar Jain et al (2019) [1], Utilized deep neural network model for facial emotion recognition and the model has been tested on two public datasets to assess the performance of the proposed model. Particularly, it has been found that the combination of FCN and residual block cloud considerably improve the overall result, which verified the efficiency of the proposed model

S. H. Wang et al (2019) [2]. Developed an improved emotion-recognition system over facial expression images. In this approach they combine the algorithm of stationary wavelet transform, single hidden layer feed forward neural network, and Jaya algorithm.

M. H. Siddiqi et al (2017) [3], proposed a real-time facial expression recognition (FER) system using smartphone camera. In order to make the system efficient and robust, in their exploration, extracted the features only from the contributing parts of the face using facial points. Once the points are extracted, then the distance and angle are calculated between these points. They utilized support vector machine (SVM). The average recognition rate for the proposed smartphone-based FER system is 85.6% for 5 basic expressions at a speed of more than 15 fps (frames per second), which is a significant improvement in real-time domain.

Alvarez et al (2018) [4], presented preliminary results on a method for facial expression detection. Such method is based on a comparison between interest points of facial images showing an emotion and facial images showing a neutral expression. The resulting gradients were classified and visualized in similarity plots that will be latter processed by a classifier. The ultimate goal is to have a real-time facial emotion recognition system of simple structure demanding little computational burden.

N. Zeng et al (2018) [5], presented a novel approach for facial expression recognition using deep sparse autoencoders (DSAE), which can automatically distinguish the expressions with high accuracy. Both the facial geometric and appearance features have been introduced to compose a high-dimensional feature with accurate and comprehensive information of emotions. Particularly, the DSAE-based deep learning framework has been established for facial expression

recognition to identify the expressions successfully by learning robust and discriminative features from data. Experiment results have demonstrated that the proposed approach outperforms the other three state-of-the-art methods in term of the accuracy rate for the 7-class expressions recognition, and provides a satisfactory performance for the 8-class expressions recognition.

Wasim Ahmed Syed et al (2018)[6], Detecting the expression associated with derision the predictor is consistently poor. Contempt, poor pre-training labelling of data, and the intrinsic difficulty at identifying contempt is clearly exhibited by the combination of lacking training and test images. In predicting emotions for test data that have expressions that do not clearly belong exclusively to one of the seven basic expressions the classifier is also not successful

S. N.	**Paper Name & Author**	**Year of Publication**	**Data Set**	**Face Detection Method**	**Feature Extraction Model**	**Expression Classification Model**	**Expressions**	**Accuracy**
[1]	"Extended deep neural network for facial emotion recognition", Author: Deepak Kumar Jain , Pourya Shamsolmoali, Paramjit Sehdev.	**2019**	Public dataset		geometric based and appearance approaches	**DNN**	Happy, sad, normal	**89%**
[2]	"Intelligent facial emotion recognition based on stationary wavelet entropy and Jaya algorithm", Author: S. H. Wang, P. Phillips, Z. C. Dong.	**2018**		Define Distance between Two Centers of Both Eyes	Wavelet Entropy	Jaya Algorithm	Happy, Sad, Disgust, Fear, Surprise, Anger, Neutral	96.8%
[3]	"Real Time Human Facial Expression Recognition System using Smartphone", Author: M. H. Siddiqi,	**2017**	Real-time	Luxand Face Recognition	Angle And Distance Method	SVM	Sad, Happy, Anger, Surprise, Neutral	85.6 %

	M. Alruwaili, J. Bang, and S. Lee.							
[4]	"A Method for Facial Emotion Recognition Based on Interest Points", Author: Alvarez, V. M., Velazquez, R., Gutierrez, S., & Enriquez-Zarate	**2018**	JAFFE	Hog	Face Alignment with Regression Tree	Multilayer Perceptron	anger, disgust, neutral, Happy, sad, surprise, fear,	88.03 %
[5]	"Facial expression recognition via learning deep sparse autoencoders", Author: N. Zeng, H. Zhang, B. Song, W. Liu, Y. Li, and A. M. Dobaie	**2018**	CK+ JAFFE	AAM, Hog, PCA	DSAE	Softmax Classifier	Happy, Sadness, Surprise, Anger, Contempt, Disgust, Fear, Neutral	95.79 %
[6]	"Emotion Detection through Facial Feature Recognition", Author: Wasim Ahmed Syed, Arun Kumar Uppalapti, Amruth Vadakattu.	**2018**	CK+	Viola-Jones's Haar-like feature cascade detector	PCA, Fisher Face + HOG	SVM	Happy, Sadness, Surprise, Anger, Disgust, Fear, Neutral	81%
[7]	"A Deep Learning Approach for Real Time Facial Emotion Recognition", Rupali Gill, Jaiteg Singh.	**2021**	Labeled Faces in the Wild	Landmark based technique	comparison of facial image from user/participant	CNN	happy, sad, disgust, surprise, fear, neutral.	93%
[8]	"Facial Emotion Detection Using Neural Network", Md. Forhad Ali, Mehenag	**2020**		Viola-Jones cascade object	HOG	SVM	Anger, Contempt, Disgust,	97%

	Khatun, Nakib Aman Turzo			detectors and Harris corner key-points			Fear, Happy, Sad, Surprise	
[9]	"Detection of Emotional States from Action Units of Facial Expressions", Abhishek Singh Kilak and Namita Mittal	**2015**	CK		Facial action coding system		Neutral, Surprise, Joy and Disgust	

3. PRESENT STUDY:

In recent year papers, for preprocessing (Face detection) mostly the Haar classifier, Adaptive skin color, Adaboost contour point algorithm and histogram equalization method is used.

For feature extraction, Gabor filter, LBP, HOG, LDA, Fisher face method etc. are used. In feature extraction, the majority of the methods are based on the texture descriptor such as LBP (Local Binary Patterns) based which gives improved results.

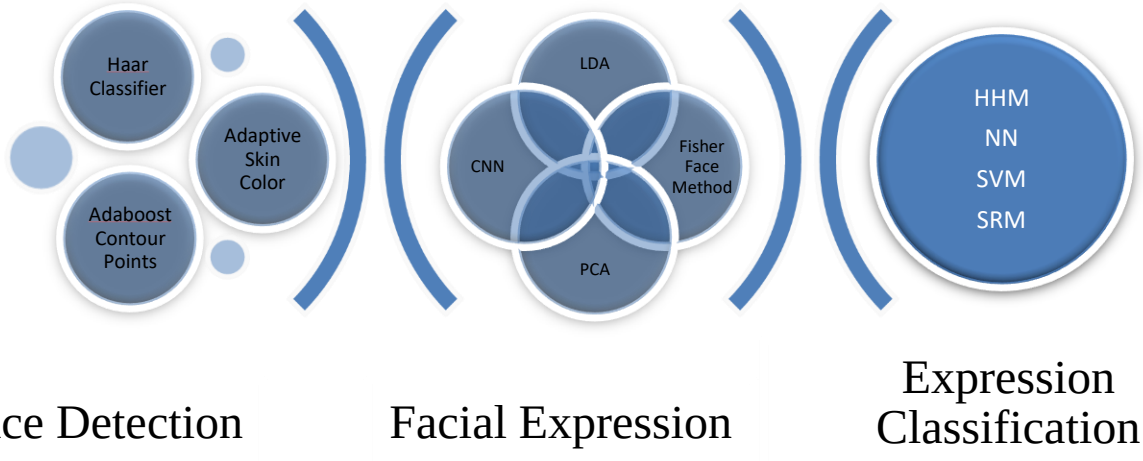


Fig. 2. Present study of FER

In modern years, the Expression classification uses the classifiers are SVM, Neural network based classifiers and pair wise classifiers, hidden markov method (HHM). The SVM classifier is highly used classifier in FER and it uses one- to –one, one- to all classification approach. Also, SVM with RBF kernel is most probably used which gives the highest classification performance comparing to other classifiers.

In 3D FER, the preprocessing of face images are performed by using the various methods such as smoothing, cropping, face alignment. The facial expressions are recognized from videos using the head gesticulation. The features such as geometric features and appearance features are

extracted from 3D faces using the various descriptors likely 3D surface descriptors, texture filters, and covariance region descriptors.

3.1 Database description

Experiments are performed on FER by using various databases likely-

The Cohn – Kanade CK-facial database [55] is the most reliable and is very widely used for research in AFEA. In this dataset, one disadvantage is that once the image sequence ends, the facial expression approaches its apex phase. So, it is inappropriate for chronological investigation. Moreover, the different record includes the date and time stamp logged on the chin of the images.

The Face Recognition Technology-FERET [56] has created its database that contains 14126 face images of over 1199 various individuals and has variations in expression, lighting, pose, and acquisition time. While developing the database they have maintained the same setup for each and every person and session who contributed in the creation of this database.

Other than the above-discussed databases for facial expressions, Few additional available databases are Japanese Female Facial Expression (JAFFE) [60], MMI FE database [61], Carnegie Mellon University-Pose, Illumination and Expression (CMU-PIE) database [63], The AR Face Database [63], Radboud Face database [64], etc.

4. LIMITATIONS:

Mostly Face Expression Recognition systems meet the problems like-

- **Subjectivity and cultural differences:** Facial expressions can vary across individuals and cultures. What may be considered a specific expression in one culture may be interpreted differently in another. This subjectivity makes it challenging to develop a universal facial expression recognition system that accurately detects and categorizes emotions across all populations.
- **Context and ambiguity:** Facial expressions can be influenced by various contextual factors, such as the environment, personal experiences, and social interactions. Recognizing emotions solely based on facial expressions may overlook important contextual information, leading to misinterpretation or misclassification of emotions. Ambiguous expressions, where emotions are not clearly displayed or easily discernible, can further complicate accurate recognition.
- **Limited facial cues:** While facial expressions are crucial for understanding emotions, they are not the sole indicator. Emotions are complex and multi-dimensional, involving other non-facial cues like body language, voice tone, and gestures. Relying solely on facial expressions may not capture the complete emotional state of an individual, leading to a partial understanding of their emotions.
- **Variability and individual differences:** People exhibit a wide range of facial expressions, and there can be significant variations in the same emotion expressed by different individuals. Factors such as age, gender, ethnicity, and personal characteristics can impact the way individuals display and perceive emotions. Training facial expression recognition models on diverse datasets that capture these variability is challenging but necessary to improve accuracy.
- **Limited sample size and biased data:** Developing accurate facial expression recognition models requires a large and diverse dataset for training. However, acquiring and annotating such datasets can be difficult, leading to relatively smaller sample sizes and potential biases. Biased data, for example, if the dataset is predominantly composed of specific demographics, can result in biased models that perform poorly on underrepresented groups or in real-world scenarios.

- **Real-time challenges:** Facial expression recognition systems often need to operate in real-time to be practical in various applications. Achieving real-time performance can be challenging due to the computational complexity involved in processing and analyzing high-resolution facial images or video streams. The delay introduced during processing can impact the system's responsiveness and usability in time-sensitive applications.

Facial expression recognition technology continues to evolve, and researchers are actively working to address these limitations. By considering the broader context, incorporating additional cues, and refining the models with diverse and unbiased datasets, the accuracy and applicability of facial expression recognition systems can be improved.

1. Methodology

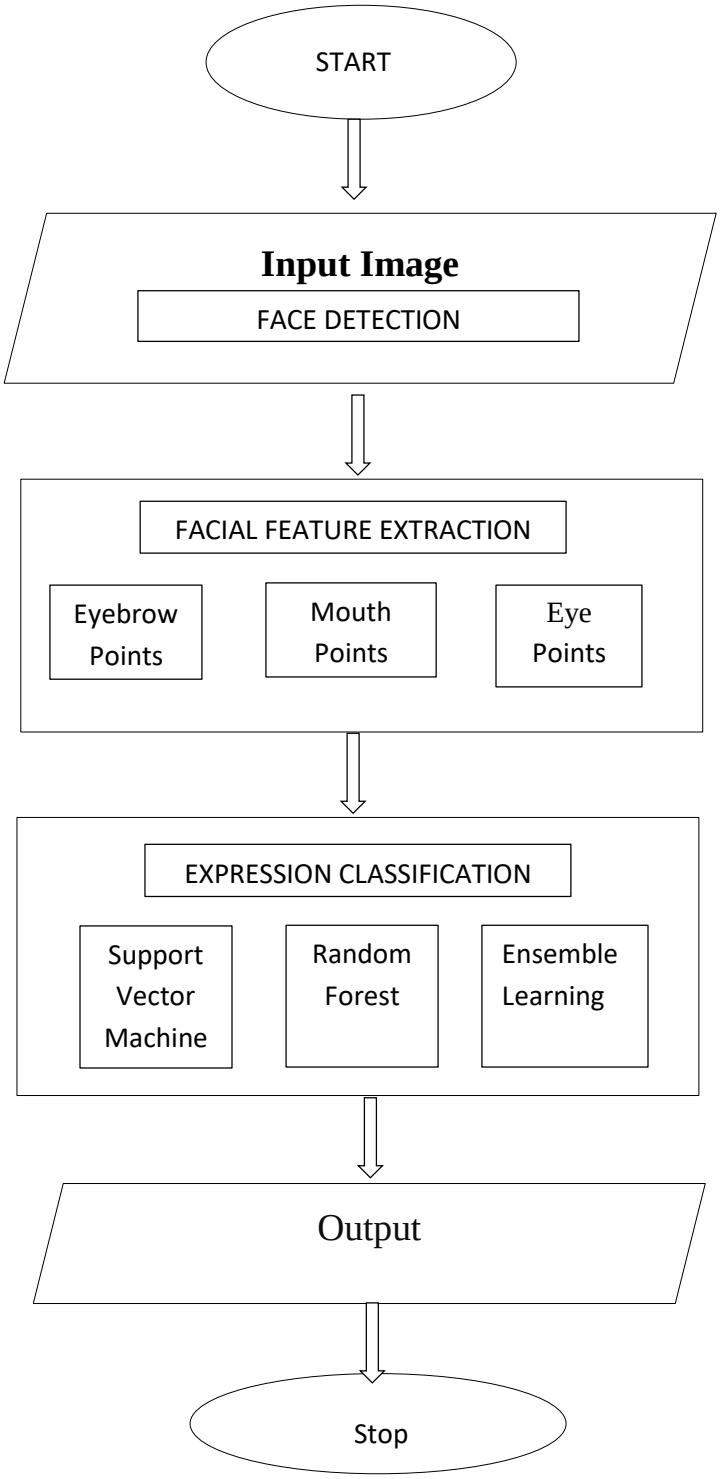


Fig. 3. Methodology used in FER

REFERENCES:

[1] Deepak Kumar Jain, Pourya Shamsolmoali, Paramjit Sehdev, “Extended deep neural network for facial emotion recognition” 2019.

[2] S. H. Wang, P. Phillips, Z. C. Dong, and Y. D. Zhang, “Intelligent facial emotion recognition based on stationary wavelet entropy and Jaya algorithm,” Neurocomputing, vol. 272, pp. 668–676, 2018.

[3] M. H. Siddiqi, M. Alruwaili, J. Bang, and S. Lee, “Real Time Human Facial Expression Recognition System using Smartphone,” vol. 17, no. 10, pp. 223–230, 2017.

[4] Alvarez, V. M., Velazquez, R., Gutierrez, S., & Enriquez-Zarate, J. A Method for Facial Emotion Recognition Based on Interest Points. 2018 International Conference on Research in Intelligent and Computing in Engineering (RICE), 1–4.

[5] N. Zeng, H. Zhang, B. Song, W. Liu, Y. Li, and A. M. Dobaie, “Facial expression recognition via learning deep sparse autoencoders,” Neurocomputing, vol. 273, pp. 643–649, 2018.

[6] Wasim Ahmed Syed, Arun Kumar Uppalapti, Amruth Vadakattu, “Emotion Detection through Facial Feature Recognition,” vol. 5, 2018.

[7] Rupali Gill, Jaiteg Singh, “A Deep Learning Approach for Real Time Facial Emotion Recognition”, December, 2021.

[8] Md. Forhad Ali, Mehenag Khatun, Nakib Aman Turzo, “Facial Emotion Detection Using Neural Network”, International Journal of Scientific & Engineering Research Volume 11, August-2020.

[9] Abhishek Singh Kilak and Namita Mittal, “Detection of Emotional States from Action Units of Facial Expressions”, International Bulletin of Mathematical Research Volume 2, Pages 281-285, March 2015.

[10] Raghav Puri, Archit Gupta, Manas Sikri, , Emotion Detection using Image Processing in Python,” in IEEE Conference ID: 42835 2018 5th International Conference, 14th - 16th March 2018

[11] Wout Swinkels, Luc Claesen, Feng Xiao, Haibin Shen, ‚SVM Point-based Real-time Emotion Detection,‘ in IEEE, 19 October 2017

[12] Chu Wang, Jiabei Zeng, Shiguang Shan, Xilin Chen, ,MULTI-TASK LEARNING OF EMOTION RECOGNITION AND FACIAL ACTION UNIT DETECTION WITH ADAPTIVELY WEIGHTS SHARING NETWORK,` in IEEE, 2019

[13] Ninad Mehendale, ,Facial emotion recognition using convolutional neural networks (FERC),` 18 February 2020

[14] G. Cao, Y. Ma, X. Meng, Y. Gao and M. Meng, "Emotion Recognition Based On CNN," 2019 Chinese Control Conference (CCC), Guangzhou, China, 2019, pp. 8627-8630.doi: 10.23919/ChiCC.2019.8866540

[15] G. Yang, J. S Saumell, J Sannie, " Emotion Recognition using Deep Neural Network with Vectorized Facial Feature" 2018 IEEE International Conference on Electro/Information Technology (EIT)

[16] Lu Lingling liu, "Human Face Expression Recognition Based on Deep Learning-Deep Convolutional Neural Network", 2019 International Conference on Smart Grid and Electrical Automation (ICSGEA)